Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 7.056

IJCSMC, Vol. 14, Issue. 11, November 2025, pg.1 – 13

An Efficient Hate Speech Detection through an Elman Recurrent Neural Network Optimized Based on Elephant Herding Optimization with a Local Search Algorithm

[1]K. Senthilkumar; [2]S. Sathiyabama; [3]D. Ayyamuthukumar

[1]Department of Computer Science, Thiruvalluvar Government Arts College, Rasipuram, Namakkal, India
[2]Department of Computer Science, Government Arts and Science College, Kangeyam, India
[3]Department of Artificial Intelligence and Data Science, Muthayammal College of Engineering, Rasipuram, India
[1] senthilkumar.kasi@gmail.com; [2] drsathyaabama@gmail.com; [3] ukdkumar@gmail.com

DOI: https://doi.org/10.47760/ijcsmc.2025.v14i11.001

Abstract: ***The rapid growth of social media sites like Twitter has heightened the spread of hate speech, and as such, there is a need for powerful and intelligent detection systems. Old-fashioned machine learning methods fail to reflect the semantic and contextual subtlety of textual information, whereas traditional deep learning networks, such as the Elman Recurrent Neural Network (ERNN), have slow convergence and are susceptible to local minima due to improper weight initialization. To address these constraints, this study will suggest the optimal model of ERNN that relies on a hybridization of Local Search Algorithm and Elephant Herding Optimization (LSA-EHO) to detect hate speech on Twitter datasets successfully. In the proposed hybrid algorithm, the global exploration power of EHO and the local exploitation power of LSA are used, which leads to a better convergence rate, solution quality, and stability. Under this setup, the hybrid LSA-EHO is an optimization of the weights and biases of the ERNN with the aim of reducing the Mean Squared Error (MSE) and increasing the classification accuracy. TF-IDF, Doc2Vec, and BERT representations of Twitter text data are preprocessed and used to index syntactic and semantic characteristics of the data. Experiment outcomes show that the ERNN-LSA-EHO model can obtain far better results compared with the current variants of optimization-based models, including ERNN-EHO, ERNN-IPSO, ERNN-PSO, ERNN-GA, and the baseline ERNN with respect to accuracy, precision, recall, F-score, and convergence performance. In particular, with BERT-based representation, the presented approach will be able to reach the accuracy of 98.12, which is significantly better than the accuracy of the next best model. The results prove that the combination of LSA and EHO increases the efficiency of search globally and locally, resulting in the highest performance of optimization and prediction.***

Keywords: ***Social media analyses, Twitter data, hate speech, Elman recurrent neural network, hyperparameters optimization***

I. INTRODUCTION

Twitter is among the numerous platforms that are inevitable in the times of social media exchange of perspectives, news, and emotions [1]. However, along with the positive interaction, hate speech is also found on these sites, and the problem with this is very socially, ethically, and psychologically bad. Hate speech recognition is a significant Natural Language Processing (NLP) issue that aims to identify and filter abusive or offensive data [2, 3]. This does not come easy due to the complexity of the social media text's informal language, its slang, emojis, and abbreviations. Traditional machine learning approaches, despite their usefulness to a certain degree, cannot depict context-specific and sequential relationships, which come into play when it comes to language, and therefore they cannot be best in terms of detection [4, 5]. Deep learning models, including Recurrent Neural Networks (RNNs) and their adaptations, show promising results since they can acquire long-term dependencies in text. One of them is the Elman Recurrent Neural Network (ERNN), and it does provide the aspects of dynamic time modeling, which is why this model can be applied to sentiment and hate speech analysis. Despite these advantages, ERNNs possess the following issues: slow convergence, local minima, and sensitivity to start weights and biases. To address these limitations, swarm intelligence-inspired optimization algorithms have been applied to optimize network parameters. However, exploration (exploring new areas) and local exploitation (refining existing solutions) are a trade-off of most generic optimization algorithms. To respond to this question, the present paper proposes a hybrid optimization approach that is based on the Elephant Herding Optimization (EHO) and Local Search Algorithm (LSA), a technique which was referred to as LSA-EHO in the optimization of the parameters of the ERNN to identify hate speech efficiently on Twitter datasets. The EHO algorithm searches the global search space using the herding behavior of the elephants, and it is among the ones that break the population into clans, and it moves position based on the best members of the clan. This will protect vast exploration and avoid an early convergence. On the other hand, the Local Search Algorithm (LSA) optimizes the optimal solutions through its neighborhood areas through refinements. The LSA-EHO hybrid approach, with a solution between exploration and exploitation, provides a faster convergence rate, accuracy of solution, and classification accuracy; it results in a combination of these two algorithms. The primary goal of the research is to train a simplified ERNN network using a hybrid LSA-EHO algorithm to enhance the effectiveness and accuracy of hate speech on Twitter. The hybrid optimization scheme would then be capable of optimizing the hyperparameters and the weight-initializations of the ERNN in a manner that it can be brought to converge within a shorter amount of time and that it can also generalize better.

The contributions of the paper are as follows,

- The new hybrid optimization method (LSA-EHO) is suggested to train the ERNN model, which is a successful combination of the global and local search methods.
- Adaptive interaction between the exploration and exploitation of EHO and LSA, respectively, results in improved parameter tuning of ERNN and faster and more stable convergence.
- The semantic meaning of hate speech material is better understood through a comprehensive feature representation that is achieved by three sophisticated techniques, which are TF-IDF, Doc2Vec-IDF, and BERT.
- Considerable experimental evidence proves the high-quality of the proposed ERNN-LSA-EHO model in its accuracy, precision, recall, F1-score, and MSE values as opposed to the available optimization-based ERNN variants.
- The proposed model adds value to the social media monitoring systems as it provides a robust and highly accurate framework of automated hate speech detection.

The structure of this paper is the follows: in Section 2, the related works are reviewed in detail within the framework of the domain of hate speech detection and optimization-based deep learning models. Section 3 outlines the research methodology proposed, such as the architecture of the optimized ERNN model and the incorporation of the hybrid LSA-EHO optimization technique. Section 4 expounds the experimental design, findings, and the analysis of the performance of different feature extraction algorithms. Lastly, Section 5 is the summarization of the paper by recapping the significant findings and the possible future research directions.

II. RELATED WORKS

Over the past few years, identifying hate speech on social media has emerged as one of the most important subjects of research in this field because of the rising amount of user-generated content. To improve the accuracy of detection, robustness, and generalization across multiple languages and modalities, researchers have used different deep learning, optimization, and hybrid techniques. A. A. Firmino et al. [6] suggested a cross-lingual learning framework, which uses BERT and XLM-R language models as pre-trained models to identify hate speech in low-resource languages with an F1-score of 92%. On the same note, N. Badri et al. [7] designed a hybrid framework that incorporates Glove and FastText embeddings and a BiGRU classifier that performed better on an OLID dataset. Deep hybrid neural models have demonstrated favorable results in bilingual hate speech classification. As an example, S. Khan et al. [8] proposed BiCHAT, a CNNBilstmAttentionBERT-based model, which scored high on Twitter datasets in terms of precision and F1-scores. Ensemble learning has

additionally been studied to support cross-domain generalization. E. In an investigation by Mahajan et al. [9]. A bagging-stacking ensemble model based on CNN, LSTM, BiLSTM, and BiGRU networks was developed and found to increase F1-scores by 4.44% relative to a baseline across nine evaluation datasets. A. J. Keya et al. [10] proposed a hybrid model, G-BERT, a combination of BERT and GRU to detect hate speech related to social media in Bengali, which was found to be accurate by 95.56. In addition,

B. A. H. Murshed et al. [11] proposed the DEA-RNN that combines the Dolphin Echolocation Algorithm with the Elman Recurrent Neural Network (ERNN) with more than 90.45% accuracy in comparison with BiLSTM and SVM models. Multimodal and graph-based approaches have also become useful in modeling complex linguistic and contextual information. M. Ali et al. [12] reported an LSTM-GRU model with community detection based on the given algorithm with 98.14% accuracy on Twitter data. A. Chhabra et al. [13] and A. Rana et al. [14] considered multimodal schemes, which combine visual and textual data with CNNLSTM and emotion-based architectures to detect multimedia hate speech, and found that F1-scores in multimodal conditions were significantly enhanced. A CNNGlove CNN-LSTM model hybrid with GloVe textual hate speech detection was suggested by Simon et al. [15], where the F1-score on Twitter and Facebook data reached 0.92. The recent research by 2025 has focused on multilingual, multimodal, and adversarial models of learning. A. Naseeb et al. [16] used Quantized Low-Rank Adaptation (QLoRA) fine-tuning, in detecting offensive language in Roman Urdu, where the LSTM attained an accuracy of 96.2%. F. K. Saddozai et al. [17] introduced a BiLSTM-EfficientNetB1 hybrid model to be used as a multimodal Urdu-English hate speech detection, with an F1-score of over 80%. A. K. Srivastava et al. [18] and E. Hashmi et al. [19] emphasized on multilingual models based on LASER, ELMo and NorBERT embeddings to detect language hate speech with low resources and proved better generalization and transfer learning properties. E. L. T. Tchokote et al. [20] created a multimodal model that is a combination of ResNet and RoBERTa with SMOTE and PCA to train the model in a balanced manner, which increased its accuracy by 18 percent. P. Kapil et al. [21] proposed a multi-task multi-modal system which combines CLIP, UNITER and BERT to detect hateful memes, and the classification performance improves to a large extent. All in all, the literature notes that the transfer learning, hybrid deep architecture, and optimization-based frameworks have achieved a lot regarding the detection of hate speech. Nevertheless, other issues like convergence stability, hyperparameter tuning, and overfitting stand out. The majority of studies are based on only global or local optimization strategies, which can reduce the diversity in solutions and the quality of convergence.

To resolve these problems, the current paper suggests an optimized Elman Recurrent Neural Network (ERNN) optimized by a hybrid Local Search Algorithm-Elephant Herding Optimization (LSA-EHO) method. This is a hybrid system that combines global exploration and local exploitation systems to enhance detection accuracy, convergence rate, and model robustness in Twitter-based hate speech datasets.

III. RESEARCH METHODS

The following subsections discuss the research methods that are used in this research work, such as ERNN, LSA, EHO, hybrid LSA and EHO, and the proposed optimized ERNN.

A. Elman recurrent neural network

Elman (1990) introduced the ERNN [22] common method, one kind of feedback neural network, the ERNN, acquires a recurrent layer developed on the basis of the hidden layer. This layer introduces a memory capability and it is a delay operator. It maintains the stability of the network on a world scale and makes it respond to the variations in time that are dynamic in nature. Topological features of the ERNN model typically have four layers: The hidden layer is fed with the information of the input layer, whose neurons are typically linear, and then employs an activation function to modify or amplify this information. Given the input of the hidden layer, the task of the connecting layer is to provide the response of the same instance as the previous one to form a local ring structure. Since the deferred memory effect of the concerned layer in the features contained in the past data, the final value of the neural network is influenced more by the real data. The results are finally taken out as the output. ERNN is trained with the help of BPNN, but it connects the output of the hidden layer and its input automatically depending on the delay and storage capabilities of the context layer. The capability of dynamic input can be enhanced by an internal feedback mechanism. ERNN is gathered based on an input, a hidden, an output, and a recurrent layer. The sample or data of each layer is transferred by one or several neurons through a nonlinear function of the weighted sum of the input samples.

$$X_{it}(k) = \sum_{i=1}^{n} X_{it}(k-1) \qquad (1)$$

In this case, X_{it} - represents an input at time t and the number of neurons n. The input of hidden neurons is as follows:

$$net_{jt}(k) = \sum_{i=1}^{n} W_{ij} X_{it}(k-1) + \sum_{j=1}^{p} C_j r_{jt}(k) \qquad (2)$$

W_{ij}- the weights of the input and hidden layer. C_j - weight between recurrent and hidden layers. The hidden layer's output is defined as follows:

$$Z_{jt}(k) = f(net_{jk}(k) = \sum_{i=1}^{n} W_{ij} X_{it}(k-1) + \sum_{j=1}^{p} C_j R_{jt}(k) \quad (3)$$

The recurrent layer is determined as follows:

$$R_{jt}(k) = Z_{jt}(k-1) \quad (4)$$

The output layer is defined as follows,

$$Y_t(k) = f(\sum_{j=1}^{p} V_j Z_{jt}(k) \quad (5)$$

The errors of the network are designed as follows,

$$E = \sum_{k=1}^{m} (t_t - y_t)^2 \quad (6)$$

Here, t_t – target value and y_t – predicted value.

B. *Elephant herding optimization (EHO)*

The EHO approach for swarm intelligence optimization was introduced by Wang et al. in 2015 [23]. The separation and the clan update operator are the two primary operators of EHO. The population of elephants is divided into multiple clans, each of which, c_i, has a set number of elephants. The matriarch (c_i) influences the elephant's new position. The definition of the elephant is as follows:

$$x_{new,ci,j} = x_{ci,j} + a \times (x_{best,ci} - x_{ci,j}) \times r \quad (7)$$

Where, $x_{new,ci,j}$- new position, $x_{ci,j}$- old position, $x_{best,ci} - x_{ci}$ is the matriarch. $r \in [0,1]$, $a \in [0,1]$. The largest elephant can be determined by the following:

$$x_{new,ci,j} = \beta \times x_{center,ci} \quad (8)$$

$$x_{center,ci,d} = \frac{1}{n_{ci}} \times \sum_{j=1}^{n_{ci}} x_{ci,j,d} \quad (9)$$

Where $1 \le d \le D$, $n_{ci,j}$is the number of elephants, $x_{ci,j,d}$is d^{th}-dimensions of elephant, $x_{center,ci,d}$ - the center of clan. In the context of solving optimization problems, the process by which male elephants split out from their family group can be described as a separation operator. The separation operator is applied by the elephant individual in each generation that is least suited, as shown.

$$x_{worst,ci} = x(xmin_{max} + 1.rand)_{min} \quad (10)$$

Where, x_{max} is the upper bound of the individual, x_{min} is the lower bounds, $x_{worst,ci}$ is the worst individual, $rand$ - stochastic distribution between 0 and 1. Elephant herding behavior serves as the inspiration for the metaheuristic optimization method. It is a member of the class of nature-inspired optimization algorithms, which use difficult optimization problems to solve by imitating the behavior of plants, animals, or other natural occurrences. It is modeled after how herds of elephant's travel, with the more powerful elephants leading the weaker ones to better grazing areas. By taking into account individual solutions and their interactions within a population, this behavior is transferred into the optimization method. EHO balances exploitation and exploration by employing both local and global search strategies. While memory aids in exploitation by directing the search toward the most well-known solutions, herding behavior assures exploration toward favorable locations.

C. *Local search algorithm (LSA)*

Local Search Algorithm (LSA) is an optimization method, which works on the assumption of searching a subset of a solution space until it is found that near-optimal solutions to complex optimization problems are found [24]. LSA, in contrast to global optimization algorithms, makes the attempt to optimize one single candidate solution by successively searching the search space, instead of exploring an entire search space. The idea behind this process of local improvement is to keep repeating it until a termination condition, such as no more improvement of the objective value or a maximum number of iterations, is met. LSA is strong because it is able to do fine-tuning of solutions obtained using global optimization techniques. It has the potential of greatly improving the quality of the solution by conducting local searches in promising areas identified by global algorithms. This renders LSA especially useful in those cases when the space of solutions is large and complicated, and cannot be explored exhaustively. The process of local exploitation ensures that convergence is more precise and fewer chances of poor stagnation occur through LSA.

D. *LSA-EHO Algorithm*

It is the hybrid optimization model which has been developed as a result of integrating the EHO algorithm with the LSA optimizations which effectively leverages the benefits of the global exploration and the local exploitation processes. EHO has much in common with extensive exploration of the search space to prevent premature convergence by a population diversity. LSA on the other is supposed to reduce the locally optimum solutions in order to increase the accuracy and quality of the answer. The joint search allows a more balanced search strategy, leading to the improvement of the convergence rate and the overall robustness of the joint search when used in a diversity of optimization problems. The proposed hybrid model implies that LSA is invoked following each EHO run to perform small process changes on the current best solution found to date. This refinement step involves having a temporary variable ($temp$) which contains the current best solution discovered by EHO and is then passed over to LSA in order to get a better solution. LSA maximizes temp via a sequence of chance selection and evasion of three features whereby each step makes a change in the feature as per its predefined parameters. The fitness of the new candidate solution is estimated upon each change. When the fitness value is higher than that of the previous solution, temp is changed, however, temp remains unchanged

otherwise. It is a procedure involving a cyclic process with the aim of ensuring that search proceeds along optimum areas of solution space or near optimal areas at the expense of computational efficiency. This implies that the hybrid EHOLSA algorithm is superior regarding the accuracy of solution, rate of convergence and stability therefore it is most suitable in the case of complex and high-dimensional optimization.

E. Proposed optimized ERNN based on EHO

The current study aims at designing an optimal ERNN model as a hybrid model of LSA- EHO to detect hate speech in Twitter data. The LSA-EHO hybrid methodology is suggested to optimize the weights and biases of the ERNN in order to increase the accuracy of detection, convergence speed, and not to be trapped in local optima. In this model, every member of the herd of elephants is a candidate solution, which is associated with a given set of hyperparameters of ERNN. The main aim of the optimization is to reduce the gap between actual and the predicted values of the ERNN model. MSE is considered as the fitness and it is mathematically defined as:

$$MSE = \frac{1}{N}(y_i - \hat{y}_i)^2 \qquad (11)$$

where y_i - predicted value. $\hat{y}_i$ - real value. N - sample's length. To reduce the difference between the actual and predicted outputs by reducing the total MSE value. Lower MSE means that the accuracy of prediction of hate speech will be higher.

This data with the best performance of the elephant is the best hyperparameter set of the ERNN model. The normalized input data is considered as the first entry of the proposed method by the enhanced ERNN. Repeated training of the network is then done after which an optimizer (LSA-EHO) is applied to adjust the model parameters to the MSE objective function. The elephant is furthermore a search agent and modified (i.e. adjustment in the position or the weights and biases of ERNN) based on the fitness analysis. EHO component is adopted to global search of the search space and LSA to local refine the best solutions to enhance speed of convergence and no premature stagnation of solutions is realized. The process of optimization continues until some termination (e.g. a small change of MSE or a set number of steps) is met. Finally, the ERNN is optimized to obtain the final predicted outputs after the process of denormalization. The suggested LSA EHO optimized ERNN model has been tried and tested to be stable, reliable, and consistent in the detection of hate speech. The hybrid swarm intelligence algorithm demonstrates the fact that hybrid approach can generate near-optimal solutions which are superior to the classical single optimization algorithms in convergent behavior and classification.

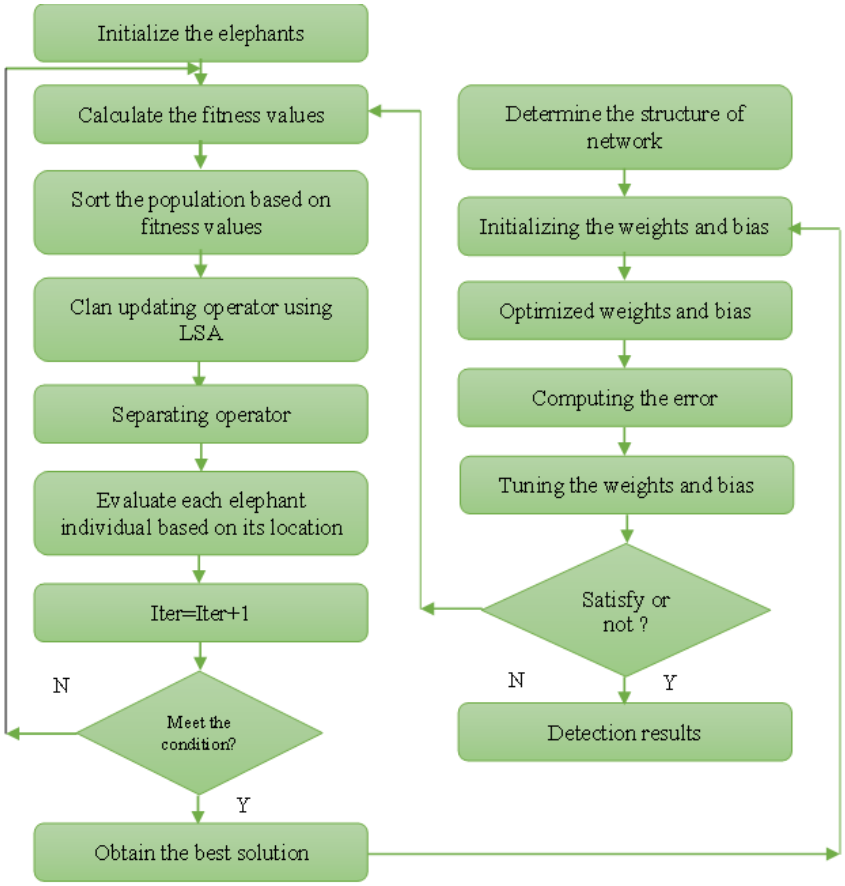

Fig. 1 : Flowchart of the proposed optimized ERNN method based on LSA-EHO

IV. Experimental Results and Analysis

The increasing rate of hate speech against anyone or a certain group of people due to the markers that define a minority on social media sites has led to the creation of sophisticated algorithms to detect hate speech. Because of the dynamism, situational and subjective quality of online communication, hate speech has not been easy to detect and control through various linguistic and situational factors. The effect of hate speech on social media platforms targeting individuals or minority groups has been growing faster, and this has necessitated the urgent development of automated methods of detection. Hate speech is difficult to identify due to complexity, contextuality and changing linguistic patterns of online communication and this makes it hard to both researchers and practitioners. To solve this problem, the current paper suggests a modified ERNN model to be improved by a hybrid LSA-EHO framework. The hybridization will help to enhance the detection accuracy, optimum convergence speed and reduction in classification error. The system aims at dividing the Twitter posts into 3 classes, which are hostile, offensive and neutral. The system is aimed at categorizing the text in social media as an offensive, hostile or neutral text. It was also necessary to extract features semantically and syntactically by using BERT, TF-IDF, and Doc2Vec techniques that offer varying degrees of contextual and statistical insights into textual information. The suggested LSA-EHO+ERNN model was compared to some of the available versions of optimization-based ERNN, such as EHO-ERNN [25], IPSO-ERNN [26], PSO-ERNN [27], GA-ERNN [28], and the ERNN [29]. The implementation of all the models in comparative terms was carried out in MATLAB 2022b on a Windows 11 platform with an Intel Core i5 processor and RAM of 8GB. The paper will discuss the datasets, parameter selection, and experimental findings, and will discuss in detail performance evaluation measures, convergence analysis, and a comparative performance of the algorithm with existing algorithms.

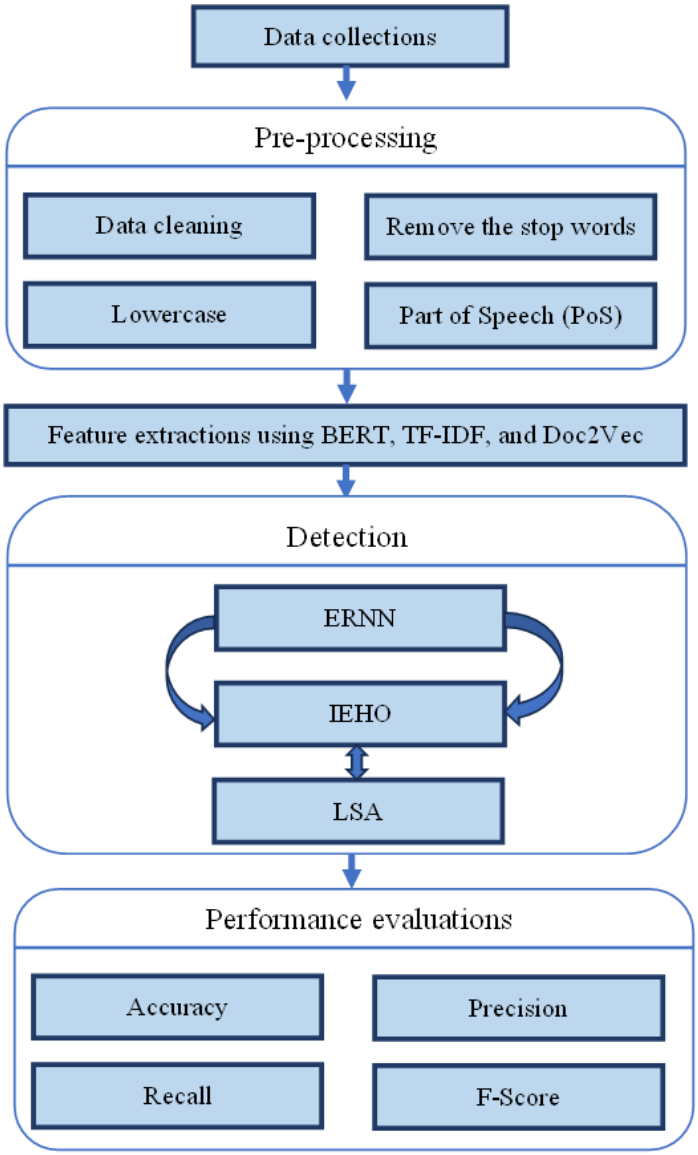

Fig. 2 : Overview of proposed hate speech detection method

TABLE I
Parameter's values

ERNN		EHO	
Parameter	Value	Parameter	Value
Activation function	Tansig	Population	10
Objective function	MSE	Clan *c*	5
Learning rate	0.05	Iteration	100
epochs	1000		
Error	0.0005		
Weight range	-0.5 and 0.5		
Activation function	Tansig		

A. Dataset

The data sets are gathered on Different social media. The dataset used in this research was found on Kaggle.com. It was prepared through a compilation of Twitter tweets. The uploaded URL lacked the description of the dataset; however, in the course of the research, it was found that the dataset only contained 31,962 tweets written in English, and no other languages were studied in the situation. 29,720 (92.98) of the tweets in the generated dataset were related to the NHS, and 2,242 (7.02) were related to high school. There is a total of ten experiments carried out during the training and testing stages to achieve the objectives of the study. Each dataset is run twenty times in order to gather statistical data regarding the precision, stability, and reproducibility of the models.

B. Data preprocessing

Data processing is an important part of the process of hate speech detection on Twitter because raw and noisy data in the form of tweets are hard to directly analyze and process; thus, using appropriate data processing, this unstructured data can be represented in a standardized form that can be readily understood and analyzed by algorithms. Data processing enhances the quality and reliability of data, eliminates irrelevant components, eliminates inconsistencies, only the most informative components, and thus simplifies computational complexity simpler and making models more efficient. The following subsection discusses the data processing as follows,

1) *Data cleaning:* The noise and unstructured nature of Twitter tweets in the form of emojis, mentions, hashtags, and URLs are a crucial component of hate speech detection that requires data cleaning. The initial one is to prepare the dataset and eliminate the missing or duplicated data to maintain uniformity. All the text can be converted to a lowercase in order to ensure the uniformity of the dataset. Secondly, the URLs, user mentions and hashtags are deleted as they seldom add to the semantic meaning of hate content. Punctuation marks, figures, and other characters are also avoided to give importance to meaningful words. Emojis may be deleted or translated to written explanations since they tend to convey some emotions that may be associated with hate speech. Short forms such as don't are extended to do not in order to maintain grammatical meaning. The repetitive characters (e.g., soooo) and slang terms are standardized to the English language.

2) *Elimination of stop words:* Stop words such as articles, prepositions, and conjunctions ought to be eliminated through regular expression matching. Stop word removal can also reduce the time taken by the system to process data, besides helping the system to minimize noise.

3) *Parts of Speech (PoS):* PoS is the identification of various words in the sentences (noun, verb, adjective, adverb, singular, plural, and remaining). This would help in establishing the negative implications of the sentence. Parts of the sentence are marked and separated, up to the point of single words.

4) *Feature extraction:* The input to the model in this work was represented as the three different feature extraction methods: BERT (Bidirectional Encoder Representations Form Transformers) — deep contextual and semantic textual relationships based on transformer-based text representations. TFIDF (Term Frequency-Inverse Document Frequency) - This is used to extract statistically relevant features, which in this case is by focusing on discriminative words in a corpus. Doc2Vec: Produces syntactically and semantically continuous dense vectors of point-sized representations of entire documents or sentences. The effectiveness of the proposed LSAEHOERNN with various textual representations was assessed with the help of these methods of feature extraction.

C. Parameter settings

The choice of algorithmic parameters would greatly determine the performance of an optimization-based learning model. To ensure a balance between exploration (searching the space of solutions worldwide) and exploitation (honing of promising areas in the search space), it is important to tune the parameters properly. The determination of the best parameter values is usually performed through empirical testing and trial and error since the performance of each parameter can be different in accordance to the problem area and the nature of the data. In the proposed LSA-EHO-ERNN model, parameter tuning has a significant role to play in improving the convergence rate, decreasing in MSE, and the overall percentage of classification. The direct impact of these

parameters on the model is that they control how the model can find optimal weights during training and avoid local minima. Elephant Herding Optimization (EHO) and the Elman Recurrent Neural Network (ERNN) have their parameters or major parameters, which are critically selected after a lot of testing to give the best results. The information about the chosen parameter settings is introduced in Table 1, which reveals the values of population size, clan update rate, number of iterations, learning rate, hidden layer units, and other suitable parameters.

D. Performance measures

The efficacy and efficiency of Swarm Intelligence (SI)-based optimization techniques are measured mostly by the performance metrics. These measures are useful in identifying the efficiency of the search and utilization of the solution space by an algorithm to find out optimal or near-optimal solutions, stable convergence behavior, and in satisfying computational efficiency and robustness in different problem domains. In this research, the proposed LSA-EHO-ERNN model and the performance of other comparative machine learning (ML) methods are evaluated based on four major key performance indicators in the form of Accuracy, Precision, Recall, and F-measure. The evaluation measures are popular in classification problems to determine the predictive power and reliability of the model [30].

1) Accuracy: The percentage of correctly recognized hate tweets over the total number of hate tweets is called accuracy. It is displayed as follows,

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (12)$$

2) Precision: It is determined as the percentage of hateful tweets determined and compared to a database of labeled hateful tweets. It could be expressed in terms of numbers as

$$Pr\,e\,cision = \frac{TP}{TP+FP} \quad (13)$$

3) Recall: It quantifies the percentage of hateful tweets that have labels that are correctly recognized and presented, which is calculated as follows,

$$Recall = \frac{TP}{TP+FN} \quad (14)$$

4) F-Score: The harmonic mean of precision and recall is the F-score that takes a value of 0 as its minimum possible value and 1 as its maximum possible value.

$$F - Score = 2 \times \frac{Pr\,ecision \times Recall}{Pr\,ecision + Recall} \quad (15)$$

True Positive (TP) of HS instances are the HS instances that happen to be classified as HS; True Negative (TN) of HS instances are the NH instances that happen to be classified as NH; False Negative (FN) of HS instances are the HS instances that happen to be classified as HS; and False Positive (FP) of HS instances are the NH instances that happen to be classified as NH. The percentage of accurate detection to total predictions is one measure a way of quantifying the accuracy of the classifier, which is simply the percentage of correct detections made by the classifier.

E. Results and discussions

The Results Analysis and Discussion section is the main interpretative section of your research paper. It is not a simple restatement of the results but expounds, interprets, and relates it to the objectives of the study, existing literature as well as the theoretical framework. It is a process that basically starts with raw numerical results and converts them to scientific knowledge. As the experimental findings denoted in Tables 2-4 and Figure 3-8 illustrate, the proposed ERNN-LSA-EHO model is effectively applied to hate speech on Twitter datasets. The multi-feature-extraction methods performance evaluation between the hybrid optimization approach and the other variants, including ERNN-EHO, ERNN-IPSO, ERNN-PSO, ERNN-GA, and the control ERNN, demonstrates good consistency in the hybrid optimization approach in terms of superior performance in comparison to the other variants.

Comparatively, features based on BERT are the most effective in terms of detection. ERNN-LSA-EHO model has a great accuracy of 98.12, and the Precision, Recall and F-score are 97.93, 98.67 and 98.29, respectively. The main reason behind this excellent performance is that BERT has a contextual embedding feature, which encodes semantic as well as syntactic associations between words. BERT, unlike other traditional methods of vectorization, can produce dynamic representations of words, which are effective in identifying subtle differences in the sense of hate and non-hate. Moreover, Enhanced Harris Hawk Optimization (EHO) and Learning-based Self-Adaptive (LSA) parameters are also integrated, which shows that the convergence rate and global search capability of the network is greatly enhanced to result in stable training and increased classification accuracy.

TABLE II
PERFORMANCE RESULTS BASED ON THE BERT METHOD

Methods	Accuracy	Recall	Precision	F-Score
ERNN-LSA-EHO	**98.12**	**98.67**	**97.93**	**98.29**
ERNN-EHO	97.36	97.02	96.74	96.88
ERNN-IPSO	95.85	95.41	94.92	95.16
ERNN-PSO	94.73	94.02	93.41	93.71
ERNN-GA	92.88	92.13	91.56	91.84
ERNN	88.91	89.46	88.72	89.09

TABLE III
PERFORMANCE RESULTS BASED ON THE TF-IDF METHOD

Methods	Accuracy	Recall	Precision	F-Score
ERNN-LSA-EHO	**96.85**	**97.56**	**96.11**	**96.82**
ERNN-EHO	95.94	95.42	93.16	94.27
ERNN-IPSO	92.99	92.01	91.79	91.89
ERNN-PSO	91.49	90.46	88.87	89.65
ERNN-GA	89.41	88.11	86.54	87.31
ERNN	85.64	86.49	84.88	85.67

TABLE IV
PERFORMANCE RESULTS BASED ON THE DOC2VEC METHOD

Methods	Accuracy	Recall	Precision	F-Score
ERNN-LSA-EHO	95.12	96.04	94.25	95.13
ERNN-EHO	93.21	93.02	91.18	92.09
ERNN-IPSO	90.86	90.24	89.80	90.02
ERNN-PSO	89.35	88.71	87.93	88.32
ERNN-GA	87.12	86.49	85.73	86.10
ERNN	83.67	83.09	82.21	82.65

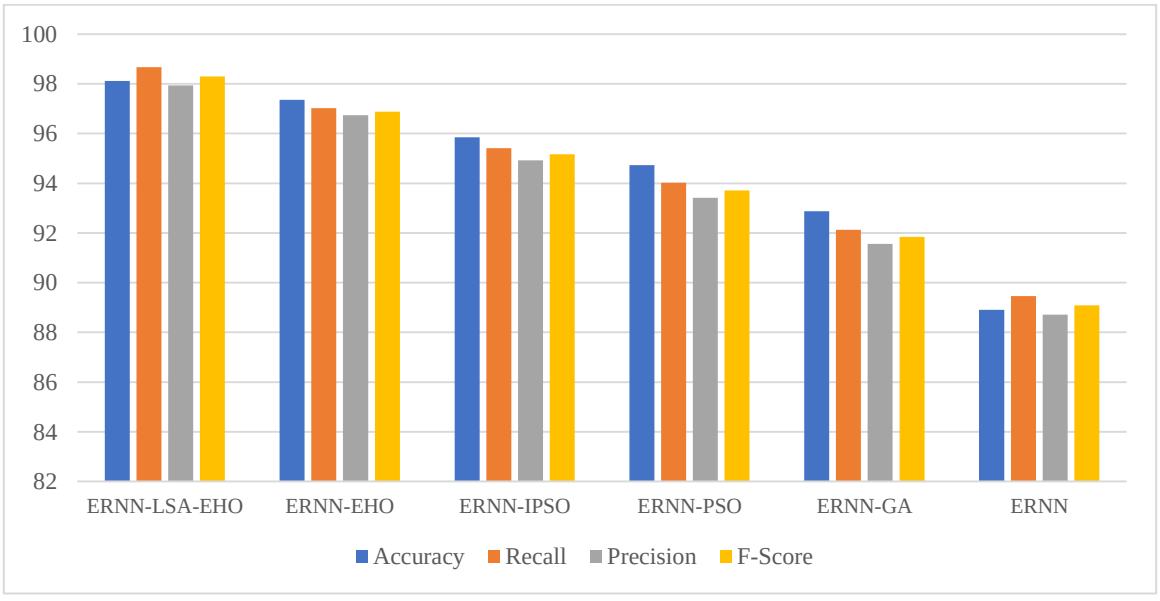

Figure 3 : Performance analysis of the detection method based on BERT

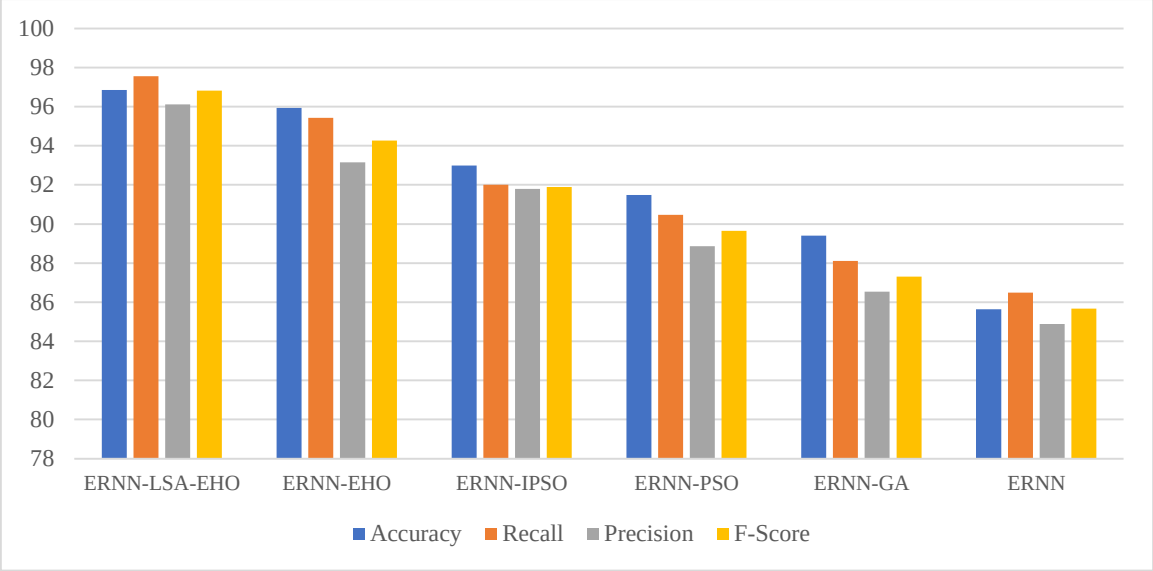

Fig. 4 : Performance analysis of the detection method based on TF-IDF

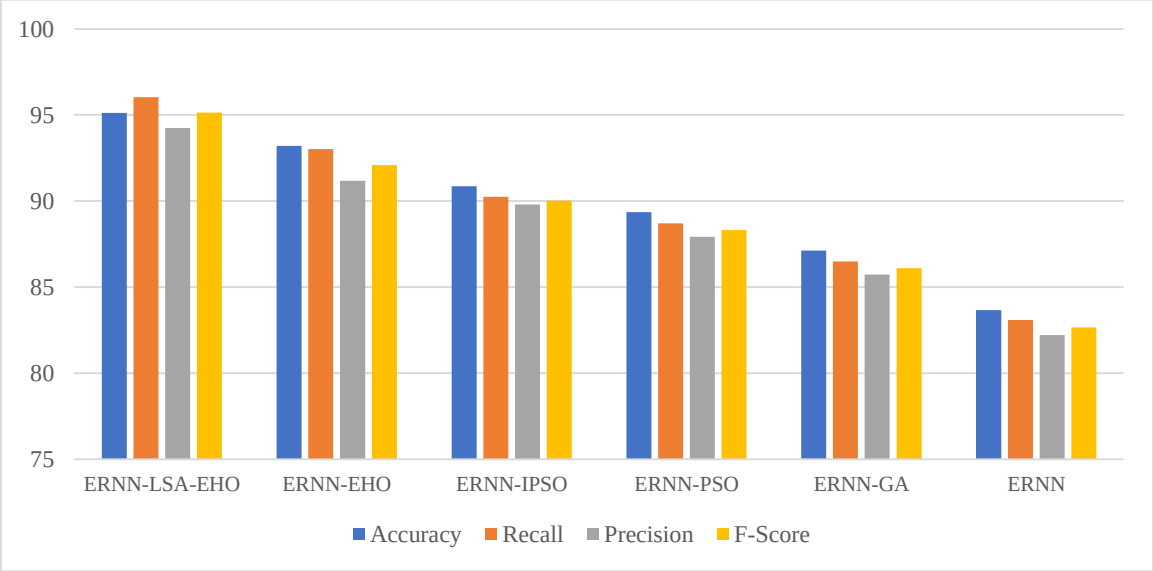

Fig. 5 : Performance analysis of the detection method based on Doc2Vec

Conversely, the models that use TF-IDF feature representations have moderate performance. Under this setup, the accuracy of the ERNN-LSA-EHO model is 96.85, which, although a bit lower than the BERT-based model, still outperforms the other baseline and hybrid ERNN models. This finding indicates the effectiveness of the optimization mechanism despite the application to more basic statistical features. But the TF-IDF method has its own fault of not having any understanding of context, as it is based on the frequency of the terms and the inversion of document frequency without having any idea of word sequence or semantically related words. This is a weakness that leads to low accuracy and recall rates, especially on ambiguous or contextual tweets.

The Doc2Vec-IDF experiments reflect equal performance of semantic representation and performance in terms of computational efficiency. ERNN-LSA-EHO model achieves 95.12 percent accuracy, which is almost 12 percent higher in comparison to the baseline ERNN. It is the interaction of Doc2Vec and IDF weighting that offers context-dependent document-level embeddings, which are meaningful and capture contextual dependencies to a smaller degree, which makes classification superior to TF-IDF. However, its results are lower than those of BERT because it lacks bidirectional contextualization and attention mechanisms based on transformers. Those findings suggest that Doc2Vec-IDF can be viewed as a potential compromise between accuracy and processing speed of resource-intensive or time-limited hate speech detectors.

The results of the convergence analysis (Figures 6-8) also support the strength of the offered approach. ERNN-LSA-EHO always shows quicker and smoother convergence of Mean Squared Error (MSE) in all the feature extraction methods. It has a learning curve stabilization rate of less than 50 iterations meaning it adapts

quickly and has very little overfitting. This is explained by the fact that the LSA strategy is self-adaptive, in that exploration and exploitation are dynamically changed as the strategy is being optimized, preventing premature convergence as it is usually the case with conventional evolutionary strategies like GA and PSO.

On the whole, the findings verify that the LSA combined with EHO optimization positively affects the learning capacity of ERNN through an effective tuning of the hyperparameters and reduced training loss. The improvement in all the methods of feature extraction indicates the generalizability and flexibility of the proposed framework. The most potent and effective model in terms of detecting hate speech is BERT + ERNN-LSA-EHO: it is able to comprehend profound linguistic undertones and contextual feelings in the text of social media, being the most precise and trustworthy one.

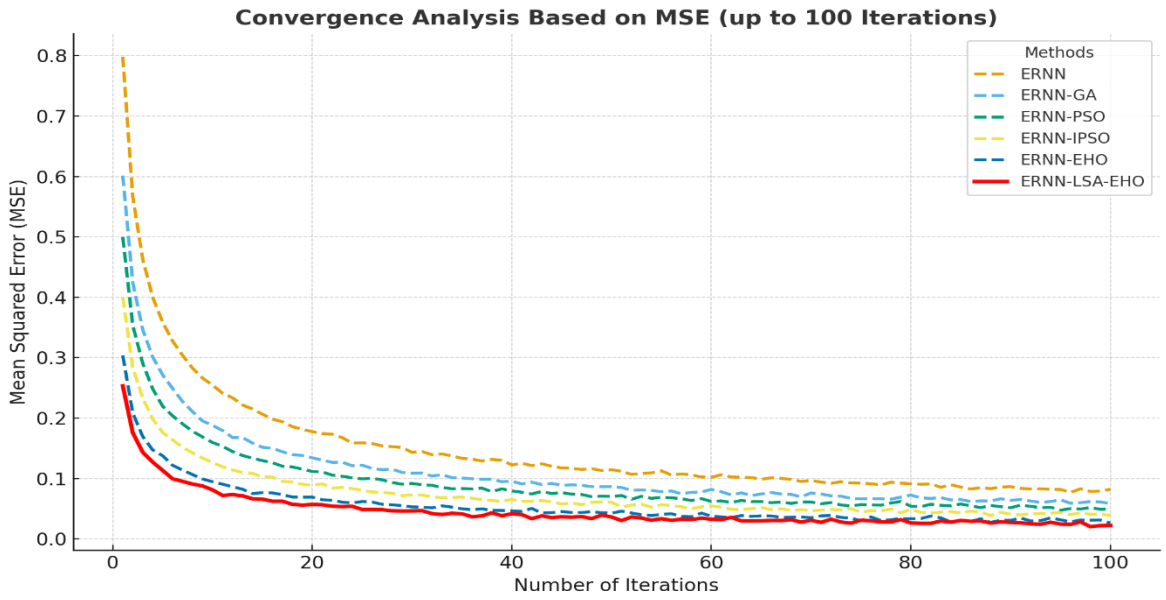

Fig. 6 : Performance analysis of detection method based on BERT

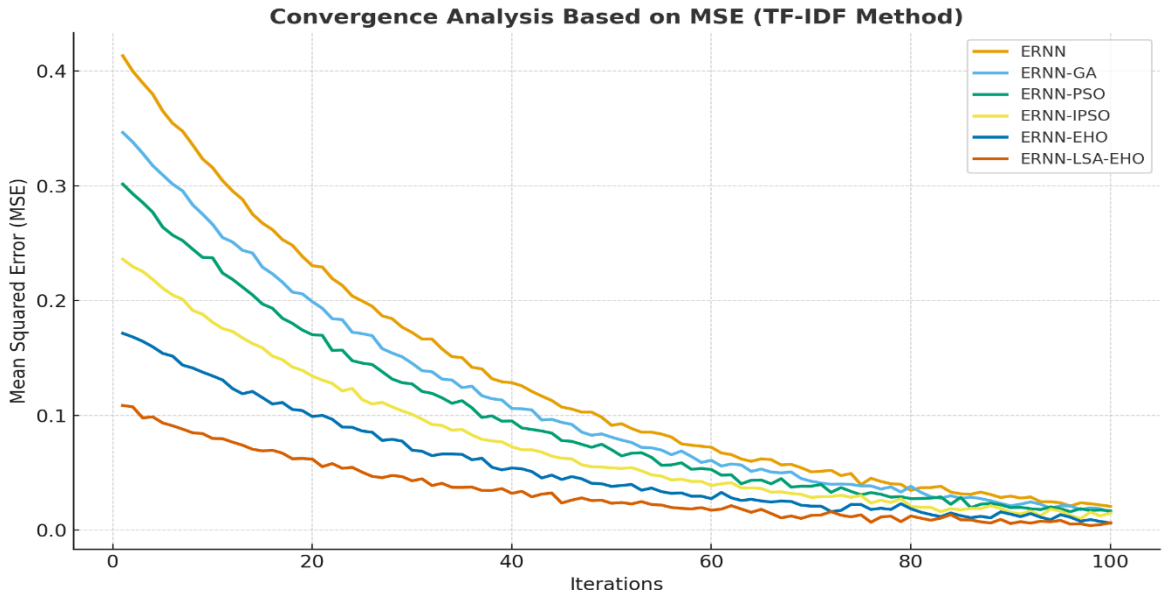

Fig. 7 : Convergence analysis of detection method based on TF-IDF

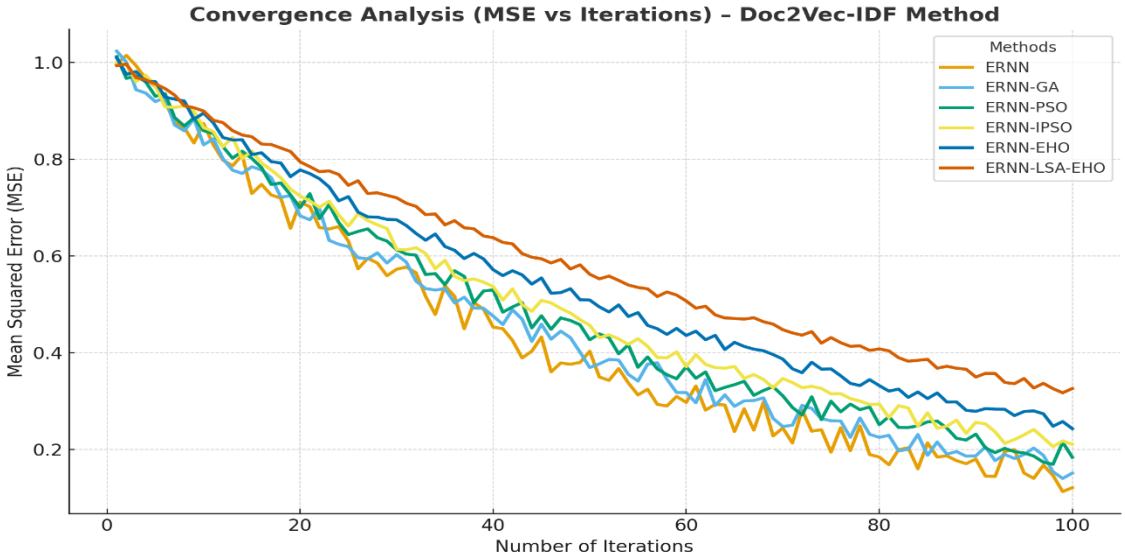

Fig. 8 : Convergence analysis of detection method based on Doc2Vec

V. CONCLUSIONS

The recommended optimized ERNN which relies on the hybrid LSA and EHO (LSA-EHO) has proven to have impressive performance in the detection of hate speech on Twitter data. The hybrid optimization method has the benefit of overcoming premature convergence, local minima issue prevalent with traditional optimization processes by combining local exploitation of LSA with the global exploration of EHO. Combination of LSA-EHO with ERNN further improves the efficiency of the training process by tuning the network weights and biases leading to faster convergence and greater accuracy of the classification. The results of the experimental analysis involving various feature extraction models - BERT, TF-IDF, and Doc2Vec-IDF are capable to prove that the suggested ERNN-LSA-EHO model is significantly more effective than the conventional and optimization-based models in the form of ERNN-EHO, ERNN-IPSO, ERNN-PSO, ERNN-GA, and the standard ERNN. In particular, the hybrid model of BERT-based embedding has the highest accuracy of 98.12, which means that it has greater semantic coverage of contextual information on hate speech detection. The results of the research confirm the fact that hybrid optimization approaches can be of great use in improving the deep-learning model performance in complex text classification tasks. The proposed LSA-EHO hybrid algorithm does not only have superior convergence properties but also provides a compromise between exploration and exploitation that is why it can be used to tasks related to optimization and learning.

The future study can be developed in various ways out of this work. To begin with, the hybrid optimization model suggested can be combined with superior deep architectures of Bidirectional LSTM, GRU, or Transformer based models to enhance understanding of context further. Second, streaming information provided by various social media can be used to monitor hate speech in real-time to test the system in terms of scalability and resilience. Also, the hybrid LSA-EHO could be customized to multi-objective optimization models to trade accuracy, computational time, and the complexity of the model. The introduction of explainable AI (XAI) methods might also contribute to the increased level of transparency and understandability of the detection process. Lastly, to make the proposed approach have wider applicability in hate speech detection systems the world over, it would be necessary to broaden the dataset to include multilingual and code-mixed tweets.

REFERENCES

[1]. Y. Zhou, Y. Yang, H. Liu, X. Liu, and N. Savage, "Deep learning based fusion approach for hate speech detection," *IEEE Access*, vol. 8, pp. 128923-128929, 2020.

[2]. J. S. Malik, H. Qiao, G. Pang, and A. van den Hengel, "Deep learning for hate speech detection: a comparative study," *International Journal of Data Science and Analytics*, pp. 1-16, 2024.

[3]. A. Albladi *et al.*, "Hate speech detection using large language models: A comprehensive review," *IEEE Access*, 2025.

[4]. M. Subramanian, V. E. Sathiskumar, G. Deepalakshmi, J. Cho, and G. Manikandan, "A survey on hate speech detection and sentiment analysis using machine learning and deep learning models," *Alexandria Engineering Journal*, vol. 80, pp. 110-121, 2023.

[5]. S. Abro, S. Shaikh, Z. H. Khand, A. Zafar, S. Khan, and G. Mujtaba, "Automatic hate speech detection using machine learning: A comparative study," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 8, 2020.

[6]. A. A. Firmino, C. de Souza Baptista, and A. C. de Paiva, "Improving hate speech detection using cross-lingual learning," *Expert Systems with Applications*, vol. 235, p. 121115, 2024.
[7]. N. Badri, F. Kboubi, and A. H. Chaibi, "Combining fasttext and glove word embedding for offensive and hate speech text detection," *Procedia Computer Science*, vol. 207, pp. 769-778, 2022.
[8]. S. Khan *et al.*, "BiCHAT: BiLSTM with deep CNN and hierarchical attention for hate speech detection," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 7, pp. 4335-4344, 2022.
[9]. E. Mahajan, H. Mahajan, and S. Kumar, "EnsMulHateCyb: Multilingual hate speech and cyberbully detection in online social media," *Expert systems with applications*, vol. 236, p. 121228, 2024.
[10]. A. J. Keya, M. M. Kabir, N. J. Shammey, M. F. Mridha, M. R. Islam, and Y. Watanobe, "G-bert: an efficient method for identifying hate speech in Bengali texts on social media," *IEEE Access*, vol. 11, pp. 79697-79709, 2023.
[11]. B. A. H. Murshed, J. Abawajy, S. Mallappa, M. A. N. Saif, and H. D. E. Al-Ariki, "DEA-RNN: A hybrid deep learning approach for cyberbullying detection in Twitter social media platform," *IEEE Access*, vol. 10, pp. 25857-25871, 2022.
[12]. M. Ali, M. Hassan, K. Kifayat, J. Y. Kim, S. Hakak, and M. K. Khan, "Social media content classification and community detection using deep learning and graph analytics," *Technological Forecasting and Social Change*, vol. 188, p. 122252, 2023.
[13]. A. Chhabra and D. K. Vishwakarma, "Multimodal hate speech detection via multi-scale visual kernels and knowledge distillation architecture," *Engineering Applications of Artificial Intelligence*, vol. 126, p. 106991, 2023.
[14]. A. Rana and S. Jha, "Emotion based hate speech detection using multimodal learning," *arXiv preprint arXiv:2202.06218*, 2022.
[15]. H. Simon, B. Y. Baha, and E. J. Garba, "A multi-platform approach using hybrid deep learning models for automatic detection of hate speech on social media," *Bima Journal of Science and Technology (2536-6041)*, vol. 6, no. 02, pp. 77-90, 2022.
[16]. A. Naseeb *et al.*, "Machine Learning-and Deep Learning-Based Multi-Model System for Hate Speech Detection on Facebook," *Algorithms*, vol. 18, no. 6, p. 331, 2025.
[17]. F. K. Saddozai, S. K. Badri, D. Alghazzawi, A. Khattak, and M. Z. Asghar, "Multimodal hate speech detection: a novel deep learning framework for multilingual text and images," *PeerJ Computer Science*, vol. 11, p. e2801, 2025.
[18]. A. K. Srivastava, M. Srivastava, S. Das, V. Jain, and T. B. Chandra, "Leveraging Deep Learning for Comprehensive Multilingual Hate Speech Detection," *Procedia Computer Science*, vol. 252, pp. 832-840, 2025.
[19]. E. Hashmi, S. Y. Yayilgan, M. M. Yamin, M. Abomhara, and M. Ullah, "Self-supervised hate speech detection in norwegian texts with lexical and semantic augmentations," *Expert Systems with Applications*, vol. 264, p. 125843, 2025.
[20]. E. L. T. Tchokote and E. F. Tagne, "Effective multimodal hate speech detection on facebook hate memes dataset using incremental PCA, SMOTE, and adversarial learning," *Machine Learning with Applications*, p. 100647, 2025.
[21]. P. Kapil and A. Ekbal, "A transformer based multi task learning approach to multimodal hate speech detection," *Natural Language Processing Journal*, vol. 11, p. 100133, 2025.
[22]. J. L. Elman, "Finding structure in time," *Cognitive science*, vol. 14, no. 2, pp. 179-211, 1990.
[23]. G.-G. Wang, S. Deb, and L. d. S. Coelho, "Elephant herding optimization," in *2015 3rd international symposium on computational and business intelligence (ISCBI)*, 2015: IEEE, pp. 1-5.
[24]. M. Tubishat, N. Idris, L. Shuib, M. A. Abushariah, and S. Mirjalili, "Improved Salp Swarm Algorithm based on opposition based learning and novel local search algorithm for feature selection," *Expert Systems with Applications*, vol. 145, p. 113122, 2020.
[25]. D. Rajakumari and D. Savitha, "An ovarian cancer prediction using an optimized Elman neural network based on elephant herding optimization," in *2023 International Conference on Emerging Research in Computational Science (ICERCS)*, 2023: IEEE, pp. 1-7.
[26]. L. Yang, F. Wang, J. Zhang, and W. Ren, "Remaining useful life prediction of ultrasonic motor based on Elman neural network with improved particle swarm optimization," *Measurement*, vol. 143, pp. 27-38, 2019.
[27]. Y. Wang, L. Wang, F. Yang, W. Di, and Q. Chang, "Advantages of direct input-to-output connections in neural networks: The Elman network for stock index forecasting," *Information Sciences*, vol. 547, pp. 1066-1079, 2021.
[28]. A. Sadeghi-Niaraki, P. Mirshafiei, M. Shakeri, and S.-M. Choi, "Short-term traffic flow prediction using the modified elman recurrent neural network optimized through a genetic algorithm," *IEEE Access*, vol. 8, pp. 217526-217540, 2020.
[29]. N. Chowdhury, "A comparative analysis of feed-forward neural network & recurrent neural network to detect intrusion," in *2008 International Conference on Electrical and Computer Engineering*, 2008: IEEE, pp. 488-492.
[30]. J. Angskun, S. Tipprasert, and T. Angskun, "Big data analytics on social networks for real-time depression detection," *Journal of Big Data*, vol. 9, no. 1, pp. 1-15, 2022.

